Stichting NIOC en de NIOC kennisbank

Stichting NIOC (www.nioc.nl) stelt zich conform zijn statuten tot doel: het realiseren van congressen over informatica onderwijs en voorts al hetgeen met een en ander rechtstreeks of zijdelings verband houdt of daartoe bevorderlijk kan zijn, alles in de ruimste zin des woords.

De stichting NIOC neemt de archivering van de resultaten van de congressen voor zijn rekening. De website www.nioc.nl ontsluit onder "Eerdere congressen" de gearchiveerde websites van eerdere congressen. De vele afzonderlijke congresbijdragen zijn opgenomen in een kennisbank die via dezelfde website onder "NIOC kennisbank" ontsloten wordt.

Op dit moment bevat de NIOC kennisbank alle bijdragen, incl. die van het laatste congres (NIOC2025, gehouden op donderdag 27 maart 2025 jl. en georganiseerd door Hogeschool Windesheim). Bij elkaar zo'n 1500 bijdragen!

We roepen je op, na het lezen van het document dat door jou is gedownload, de auteur(s) feedback te geven. Dit kan door je te registreren als gebruiker van de NIOC kennisbank. Na registratie krijg je bericht hoe in te loggen op de NIOC kennisbank.

Het eerstvolgende NIOC vindt plaats op 18 maart 2027 in Arnhem en wordt georganiseerd door HAN University of Applied Sciences.

Reacties over de NIOC kennisbank en de inhoud daarvan kun je richten aan de beheerder:

R. Smedinga kennisbank@nioc.nl.

Vermeld bij reacties jouw naam en telefoonnummer voor nader contact.

Aandacht voor gegevens

C.M.Veenstra-Strijland
Nijenrode, Universiteit voor bedrijfskunde
Straatweg 25
3621 BG Breukelen

Samenvatting

In een introductiecursus informatica dient aandacht besteed te worden aan gegevensstructurering. Normalisering is voor inleidende cursussen een te complexe methode. Deze paper beschrijft een vrij intuïtieve methode om een gegevensmodel te ontwikkelen.

1 Inleiding

Van een informatiesysteem is de gegevensbank de centrale component. Wanneer leerlingen nu in een inleidende cursus informatica voor alles zich bezig houden met het opdoen van kennis van de bouw en de werkwijze van de computer, terwijl de praktische oefeningen bestaan uit programmeren in bijvoorbeeld PASCAL, dan raakt het belang van de gegevens op de achtergrond. Zonder te abstract te worden, moeten toch ook deze leerlingen nadenken over wat gegevens zijn, hoe het onderlinge verband is en hoe dat in een computergeheugen gerealiseerd kan worden.

2 Een definitie van een informatiesysteem

Davis en Olsen (1987:7) noemen in hun definitie van een management-informatiesysteem als onderdelen: 'computerapparatuur, computerprogrammatuur, met de hand uitgevoerde procedures, modellen voor analyse, planning en beheersing, besluitvorming en een gegevensbank'. Op deze gegevensbank, in het algemeen de gegevens, wil ik mij in dit verband nader richten. Temeer daar deze auteurs als definitie van een management-informatiesysteem geven: 'een geïntegreerd mens-machine systeem waarmee informatie kan worden verkregen, die dient ter ondersteuning van de operationele activiteiten, het management en de analyse- en besluitvormingsfuncties in een organisatie'.

In wezen is de gegevensbank de basis van het systeem, de gegevens vormen de 'grondstof' waarmee het informatiesysteem het hierboven geformuleerde doel kan bereiken. Deze gegevens moeten goed beheerd worden om optimale beschikbaarheid te garanderen. In het beeld waarbij de gegevens de kern van het informatiesysteem vormen, zijn de computerapparatuur, de computerprogrammatuur, de procedures en de modellen niets anders dan hulpmiddelen waarmee het doel van het systeem bereikt kan worden.

Een volgende stap is te komen tot een definitie voor 'gegevens'. Davis en Olsen (1987:10) geven in het begin van hun boek wel een definitie van informatie: 'gegevens die betekenis hebben of nuttig zijn voor de persoon die ze ontvangt'. Opvallend is dat vóór het begrip gegeven gedefinieerd wordt, dit begrip in een andere definitie gebruikt wordt. Pas later (pag.115) wordt een nadere omschrijving gegeven van gegevens: 'symbolen die de realiteit weergeven, beschrijven of vastleggen'. Soms, in bepaalde situaties, onder bepaalde voorwaarden, via een zekere wijze van weergeven, etcetera, zal men dit gegeven informatie kunnen noemen. Dit onderscheid is niet eenvoudig, in de praktijk worden de beide begrippen dan ook door elkaar gebruikt.

3 Het ontwerp van de gegevensbank

Bij het bouwen van de gegevensbank wordt als eerste bepaald welke informatie het systeem moet opleveren, de informatiebehoeftebepaling. Vandenbulcke (1984) omschrijft dit als 'het opsporen en identificeren van relevante informatiebehoeften bij geautoriseerde gebruikers'. Het resultaat van deze activiteit is het informatiemodel. In de volgende fase, de gegevensanalyse, met als resultaat het gegevensmodel, wordt, op basis van het informatiemodel, de gegevens en de gegevensstructuren vastgelegd. Tenslotte volgen nog activiteiten die samenhangen met meer technische aspecten van het gegevensgebruik, in dit kader van minder belang.

Voor het bepalen van de informatiebehoefte (de eerste fase dus) zijn diverse strategieën mogelijk, waarbij een zeer voor de hand liggende is te kijken naar de mogelijkheden die het bestaande informatiesysteem biedt. Met het 'bestaande informatiesysteem' wordt niet persé een geautomatiseerd systeem bedoeld. Het gaat om de informatie die men op dit moment gebruikt bij operationele activiteiten of bij besluitvorming. En hoewel deze strategie een aantal gevaren in zich draagt (men zal zich teveel laten leiden door de huidige

mogelijkheden èn besteedt mogelijk te weinig aandacht aan wat juist nu niet mogelijk is), is dit zeker in onderwijssituaties een goed herkenbaar uitgangspunt. In boeken over gegevensbanken wordt deze methode dan ook veel gebruikt (bijvoorbeeld Eilers 1986). Men kan dit wel beschouwen als een manier van 'terugwerken', immers, de uitkomst is bekend, en de leerling moet op basis daarvan uitzoeken hoe de gegevensbank er uit zal moeten zien om aan de eerder beschreven voorwaarde te kunnen voldoen.

Verwarrend is nu wel (ook Eilers 1986) doet dit) dat vaak geen onderscheid wordt gemaakt tussen de informatiebehoeftebepaling en de gegevensanalyse, zij worden als één activiteit gezien. Volgens de definitie van gegevens is dit juist, immers de in- en uitvoer van het systeem zijn gegevens, didactisch ware het echter beter wel dit onderscheid te handhaven. Temeer daar de gevoerde strategie (het uitgaan van de huidige mogelijkheden van het systeem) altijd het gevaar van zich blindstaren op wat nu mogelijk is, in zich draagt.

Gaat men hier echter aan voorbij, dan zal het resultaat van de informatiebehoeftebepaling een (veelal grafische) weergave zijn van de informatie-elementen en hun onderlinge relatie.

In de volgende fase, de gegevensanalyse, ligt het zwaartepunt van de werkzaamheden op het aanbrengen van een structuur in de gegevensverzameling, wat veelal gebeurt via de normalisatie-methode van Codd. Ook hier kan men het resultaat (het gegevensmodel), zeker wanneer men dit in een onderwijssituatie doet, grafisch weergeven.

4 Analyse van enkele curricula met betrekking tot het ontwerp van de gegevensbank

Wanneer men ervan uitgaat dat de gegevensbank de centrale component van een informatiesysteem is, dan moet reeds in een introductiecursus informatica of informatiekunde hier voldoende aandacht aan besteed worden. Enkele curricula zullen hiertoe aan een nader onderzoek onderworpen worden.

Met name opleidingen waarvan de te bereiken doelen van een afzonderlijke cursus goed omschreven zijn, vormen nu het uitgangspunt. De eigenlijke eindtermen van de totale opleiding (zo deze beschreven staan), vormen geen houvast, daar het uitgangspunt de introductiecursus betreft.

Opleidingen waarvoor de beschrijving aan deze eis

voldoen, vindt men zowel binnen het regulier als het niet-regulier onderwijs, in het bijzonder echter bij het schriftelijk onderwijs. Bij het regulier onderwijs wordt meestal volstaan (zo men dit tenminste heeft gedaan) met het formuleren van de eindtermen (bijvoorbeeld Schilder 1989) van de totale opleiding.

4.1 Basisvorming

De voorlopige eindtermen voor de basisvorming voortgezet onderwijs betreffen (zoals oorspronkelijk vastgesteld) slechts een twintiguur-vak, zodat het weinig zinvol is hier verder aandacht aan te besteden. In het advies dat de ontwikkelgroep onder leiding van prof.Plomp heeft uitgebracht over de voorlopige eindtermen (1988) (waarbij de commissie uitgaat van een vak van 120 uur) ontbreekt als doelstelling het kunnen 'vertalen' van de informatiebehoefte in een gegevensverzameling echter geheel.

4.2 AMBI-PDI

Bij de AMBI-modules wordt (na de cursussen HE.0 en HE.1) pas in de cursus Elementaire Informatiekunde (HE.2) enige (zeer minieme) aandacht besteed aan gegevensanalyse. Pas in de cursus Gegevensanalyse, -modellering en -beheer (HB.2) komt het onderwerp diepgaand aan de orde (Stichting EXIN 1989).

Ook in de opleiding voor de PDI-modules wordt vrijwel geen tijd besteed aan gegevensmodellering en -structurering. In Stichting EXIN (1991), dat een beschrijving geeft van de examenspecificaties per module, is het onderwerp pas terug te vinden bij de module Programmeerprincipes (MP.1). Hier staan de normalisatiestappen tot en met de derde normaalvorm vermeld. Gezien het geringe aantal studiebelastingsuren (10), waarbij nog heel wat andere onderwerpen aan de orde moeten komen, zal dit echter slechts uiterst globaal kunnen gebeuren.

Naast de beschrijving van curricula geven ook leerboeken een indicatie. Een voorbeeld is Pollaert (1988), te gebruiken voor de eerste cursus van de PDI-opleiding, Technische middelen, waarin geen aandacht aan de informatiebehoefte wordt besteed. Ook in een boek voor de tweede cursus, Systemen en toepassingen, waar toch tenminste het belang van dit onderwerp genoemd had moeten worden, is het onderwerp niet te vinden (Papo 1988).

4.3 Open universiteit

In de introductiecursus informatica van de Open

universiteit wordt 25% van de tijd besteed aan het blok gegevensanalyse, met als ondertitel 'van informatiebehoefte naar gegevensstructuur'. De gehanteerde methode is de normalisatietechniek.

4.4 Conclusie

Hoewel hier in het bovenstaande niet verder op is ingegaan, wordt in de meeste cursussen (en ook de betreffende studieboeken) wel uitgebreid aandacht besteed aan de componenten apparatuur en programmatuur. Onder het hoofdstuk gegevens krijgt de bestandsorganisatie (sequentieel, direct, etc.) veel aandacht, maar gegevensstructurering blijkt dus vaak het zorgenkind. Bij praktisch computergebruik staat programmeren vaak centraal. Voor de leerling zal het belang van de gegevenscomponent mogelijk niet voldoende duidelijk worden.

5 Gegevensstructurering via een intuïtieve methode

In het voorgaande is gesteld dat gegevens, en dan in het bijzonder gegevensstructurering, te weinig aandacht krijgen in een introductiecursus informatica. Een mogelijke reden hiervoor is dat het gegevensmodel gewoonlijk verkregen wordt via het proces van normaliseren, hetgeen voor beginnende leerlingen veelal als te moeilijk wordt beschouwd.

Wanneer nu toch aandacht besteed wordt aan gegevensstructurering, dient voor een meer eenvoudige methode gekozen te worden om dit doel te bereiken. In deze paragraaf staat beschreven de manier waarop dit op Nijenrode gebeurt, in de eerstejaarscursus informatiekunde.

Hierbij wordt uitgegaan van een informatiebehoefte die de studenten bekend overkomt. Het voordeel van het kiezen van een dergelijke informatiebehoefte, dus uit de belevenissfeer van de studenten, is dat de studenten eerder geïnteresseerd zijn in het oplossen van het probleem.

```
                                   A.B. Leergraag
                                   Appelgaarde
                                   1234 AX Nieuwstad

Breukelen, 04-07-1989


Bij de controle van uw kamer nummer:   11
zijn de volgende schades geconstateerd:

omschrijving:                    bedrag:

IJsbakje uit koelkast weg          2.50
Toegangsdeur beschadigd           20.50
                                  -----

Totaal                            23.00

Factuurnummer:   43
```

Figuur 1
Een factuur

Toelichting: op Nijenrode wonen de studenten gedurende de eerste twee jaar van hun studie verplicht intern op de campus. De kamers worden door Nijenrode aan de studenten verhuurd. Aan het begin van de zomervakantie worden de kamers ontruimd en vindt een controle plaats van de status van de kamers. Deze controle vindt plaats met een handterminal waarbij per kamer een lijst afgewerkt wordt. Van eventuele schades wordt het nummer, samen met het kamernummer, ingetikt. Bij elk schadenummer hoort een bedrag. De student ontvangt een rekening van de schade aan de gehuurde kamer.

Op basis van dit document kan de student nu, stap voor stap, gestuurd door de docent, een aantal vragen gaan beantwoorden. Hierbij kunnen in de loop van de oefening tevens een aantal begrippen behandeld worden. Bijvoorbeeld:

1 Wat zijn in dit document de relevante gegevens die in je in een database moet opslaan om later deze factuur te kunnen maken? De vraag luidt dus zoiets als: 'waar gaat het hier om?'. Of wel, 'omcirkel alles wat van belang is'.

Het is nu belangrijk om volledig te zijn, maar, anderzijds, om gegevens die niet in de database hoeven te worden opgeslagen (zoals de begeleidende tekst) nu al weg te laten.

A.B. Leergraag
Appelgaarde
1234 AX Nieuwstad

Breukelen, 04-07-1989

Bij de controle van uw kamer nummer: 11
zijn de volgende schades geconstateerd:

omschrijving:	bedrag:
IJsbakje uit koelkast weg	2.50
Toegangsdeur beschadigd	20.50

Totaal	23.00

Factuurnummer: 43

Figuur 2
Een factuur met relevante gegevens omcirkeld

2 Hoe kunnen deze gegevens in groepen gerangschikt worden? Benoem daarbij elke groep.
Nu kan tevens het onderscheid tussen gegeven (object) en bijbehorende kenmerken (attribuut) gemaakt worden.

De student zal waarschijnlijk wel de objecten student, schade en factuur onderscheiden, elk met bijbehorende attributen.

3 Zijn er nog attributen die gesplitst moeten worden in afzonderlijke attributen?
Dit levert het begrip enkelvoudig gegeven. Tevens kan besproken worden dat er soms toch iets voor te zeggen valt dat een attribuut wel uit meer dan één gegeven bestaat.

In dit voorbeeld kan men discussiëren over het al of niet splitsen van de postcode en het adres, evenzo over het al of niet splitsen van de voorletters en de achternaam.

4 Zijn er gegevens die misschien niet opgeslagen hoeven te worden omdat ze berekend kunnen worden.
Dit betreft de zogenaamde procesgegevens.

Hier betreft dit het totaal.

5 Kan op basis van de gegroepeerde gegevens aan de informatiebehoefte, de factuur dus, nog wel voldaan

worden?
Nu komt men (met wat hulp van de docent) tot de definitieve groepering van de gegevens.

Meestal zal de student de vorige stappen wel kunnen uitvoeren. Immers, er zijn nu gegevens over studenten, over schades en over facturen. Maar het probleem ligt bij het kunnen combineren dat bij een bepaalde factuur bepaalde schades horen. De student zal nu de gedachtensprong naar factuurregels moeten maken, met als attributen het factuurnummer en de omschrijving van de schade.

6 Hoe kan men binnen een grote verzameling van gegevens die uit veel regels waarden bestaan één regel aanwijzen?
Het begrip sleutel plus de eis van een unieke waarde volgt hieruit. Zeker een sleutel bestaande uit meerdere attributen, levert weer een kleine hindernis op.

7 Hoe schrijft men dit op?
Voor de schrijfwijze van de groepen valt aan te raden de standaardschrijfwijze van de relationele tabeldefinitie te gebruiken. Om het verband aan te geven tussen de objecten onderling kan een Bachmandiagram gebruikt worden, maar in plaats van beide schrijfwijzen kunnen de resultaten ook bijvoorbeeld in een bubble chart of een entiteitassociatiediagram (Infomod) worden vastgelegd.

6 Theoretische ondersteuning

In de loop van het bovenstaande proces kan de docent, op een wijze aangepast aan het kennisniveau van de leerling een aantal meer fundamentele begrippen behandelen. Juist omdat het voorbeeld de leerling bekend voorkomt, zal het belang van deze zaken ook eerder ingezien worden.

Voorbeelden zijn:
-het bestaan van verschillende informatiebehoeften binnen één organisatie: niet overal is men in dezelfde gegevens van één leerling geïnteresseerd, maar ook over welke leerlingen men iets wil weten, kan variëren;
-problemen bij het dubbel opslaan van hetzelfde gegeven;
-gegevensonafhankelijkheid;
-privacy- en beveiligingsaspecten.

7 Het gebruik van de gegevensbank

Optimaal is de situatie wanneer de leerling, na enig inzicht verkregen te hebben in de vertaling van de informatiebehoefte naar een gegevensmodel, ook via een computerpracticum deze nieuw verkregen inzichten kan toepassen. Het wordt voor de leerling mogelijk om te controleren of de oorspronkelijke informatiebehoefte ook werkelijk te verwezenlijken valt met de gegevensbank, zoals gerealiseerd in de computer.

Voorwaarden hiervoor zijn dat:
1 De opslagwijze zo direct mogelijk aansluit bij de manier waarop de gegevens in het gegevensmodel zijn weergegeven;
2 De taal waarmee de gegevensbank wordt opgebouwd en de taal waarmee de gegevensbank wordt geraadpleegd (bij voorkeur dient hiervoor één identieke taal gebruikt te worden) eenvoudig aan te leren zijn;
3 De taal een voldoende aantal functies kent voor het raadplegen van de gegevensbank.

Het ligt voor de hand, uitgaande van de bovenstaande voorwaarden, te kiezen voor een relationele gegevensbank en de taal SQL. Diverse relationele pakketten leveren tegenwoordig hulpmiddelen voor de eindgebruiker waarmee menu-gestuurd met invulschema's een keuze gemaakt kan worden uit de mogelijkheden.

Bij het gebruik van SQL beschikt de leerling gelijk over alle mogelijkheden van een relationele taal. Daarnaast ziet men echter pogingen programmatuur te ontwikkelen die meer de mogelijkheid van een didactisch pad kent. Van het werken met een bestand (één tabel) werkt de leerling via diverse niveaus van abstractie toe naar het relationele model. Een voorbeeld hiervan beschrijft Bergervoet (1990). Hoewel hier nog geen resultaten over bekend zijn, is mogelijk deze methode aan te raden bij leerlingen met weinig computerervaring, dus bijvoorbeeld in het voortgezet onderwijs.

8 Afsluiting

Het doel van de bovenstaande aanpak is voor alles dat de leerling inzicht krijgt in het belang van de gegevenscomponent van het informatiesysteem. Wanneer de oefening in gegevensstructurering gevolgd wordt door een computerpracticum geldt nog wel steeds dat de daarvoor te gebruiken programmatuur nu, in 1990, redelijk veel kennis vereist, waardoor toch weer te veel nadruk op de techniek moet worden gelegd. Misschien dat in de toekomst, wanneer het interface

tussen computer en gebruiker echt 'vriendelijk' gaat worden, dit probleem op de achtergrond gaat raken!

Gebruikte literatuur

Bergervoet, P.M.A. (1990) De Databaas, Towards a Didactics of teaching relational Databases in Secondary School. Nog niet gepubliceerde paper ten behoeve van de 5e World Conference on Computers and Education, juli 1990.

Davis, G.B. en M.H. Olsen (1987) Management informatiesystemen. Schoonhoven: Academic Service.

Eilers, H.B., W.F. Jansen en H.H.J.de Volder (1986) SQL in de praktijk. Schoonhoven: Academic Service.

Mulder, F. en B.W. Weber-Meijerhof (1988) Introductiecursus informatica. Heerlen: Open universiteit.

Papo, M.A. (1988) Systemen en toepassingen. Schoonhoven: Academic Service.

Plomp, Tj. e.a. (1989) Advies voorlopige eindtermen basisvorming voortgezet onderwijs. Enschede.

Pollaert, W. en R. Jacobs (1988) De technische middelen. Schoonhoven: Academic Service.

Schilder, J. (1989) Curriculum informatiekunde voor het HEAO. Enschede: HOI.

Stichting EXIN (1989) AMBI 88; EXAMENPLAN 1990. Maarssen: Kluwer/Examenpublikaties AMBI.

Stichting EXIN (1991) PDI; examenplan. Maarssen: Kluwer/Examenpublikaties AMBI.

Vandenbulcke, J.A. (1984) Data base systemen voor de praktijk. Antwerpen/Deventer: Kluwer.

WRR-rapport (1988) Basisvorming in het onderwijs. Zoetermeer: min.O&W.